# PATENT ABSTRACTS OF JAPAN

(11)Publication number :

01-257999

(43)Date of publication of application: 16 10.1989

(51)Int.Cl

G10L 9/14

(21)Application number: 63-085191

(71)Applicant:

**NEC CORP** 

(22)Date of filing:

08.04.1988

(72)Inventor:

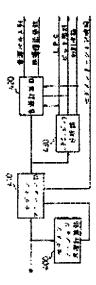
OZAWA KAZUNORI

# (54) VOICE SIGNAL ENCODING AND DECODING METHOD. VOICE SIGNAL ENCODER AND VOICE SIGNAL DECODER

#### (57)Abstract:

PURPOSE: To reduce the amount of information required for sound source signal transmission by segment a voice signal nonuniformly by using feature parameters, and finding a sound source pulse train with all the whole section except at a vowel stationary part.

CONSTITUTION: The voice signal is divided nonuniformly by using the feature parameters and when a divided section is a vowel stationary part which has a little variation in the features of a voice and is long in time, a sound source pulse train is found for one pitch sections among the sections. but when the section is not the vowel stationary part the sound source pulse train is found with all the sections. Namely, the discrete voice signal is inputted on a transmission side and a segmentation part 410 divides the voice signal into nonuniform sections and sound source signals of all or some of the divided sections are represented as a combination of pulse trains calculated by a sound source calculation part 420 and transmitted. Then, those pulse trains are used on a reception side to restore the sound source signals of said sections, and a composite voice signal representing the voice signals excellently is outputted. Consequently, the amount of information required for the sound source signal transmission is reduced



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiners decision of rejection]

[Date of requesting appeal against examiner's decision of rejection

[Date of extinction of right]

⑩日本国特許庁(JP) ⑪特許出願公開

# @ 公開特許公報(A) 平1-257999

®Int. Cl. ⁴

識別記号

庁内整理番号

❸公開 平成1年(1989)10月16日

G 10 L 9/14

J -8622-5D

審査請求 未請求 請求項の数 3 (全8頁)

◎発明の名称

音声信号符号化復号化方法並びに音声信号符号化装置及び音声信号 復号化装置

②特 頭 昭63-85191

②出 願 昭53(1988)4月8日

小澤 一範 @発明 者

東京都港区芝5丁目33番1号 日本電気株式会社内

日本電気株式会社 の出 願 人

東京都港区芝5丁目33番1号

個代 理 人 弁理士 岩佐 義幸

## 明報書

#### 1. 発頭の名称

音声信号符号化復号化方法並びに音声信 母符号化装置及び 資声信号復号化装置

#### 2. 特許請求の範囲

(!) 離散的な音声は号を入力し聴覚の特性と対 恋の良い方法により前記音声信号を非一様な区間 に分割し、その分割された区間の全部または一部 の区間における音波信号を複数個のパルス列の組 合せで表して伝送し、

伝送されたバルス列を用いて前記区間の音楽信 号を復元して前記音声信号を変す合成音声信号を 出力する音声信号符号化復号化方法。

(2) 入力した難放的な音声信号系列から聴覚の 特性と対応の良い特徴パラメータを抽出しそのパ ラメータを用いて音声信号を非一様な時間区間に セグメンテーションするセグメンテーション回路

分割された音声信号の全部または一部の区間か ら短時間スペクトル特性を表すスペクトルパラメ

ータとピッチパラメータとを計算するスペクトル パラメータ計算回路と。

分割された区間の全部または一部の区間におけ る音楽信号を表す複数個パルス列の組合せを計算 する音源パルス計算風路と、

スペクトルパラメータとピッチパラメータと音 源パルス列を組み合わせて出力するマルチプレク サ回路とを有する音声は号符号化装置。

(3) 音声信号の短時間スペクトル特性を表すス マクトルパラメータとピッチパラメ…タと音源信 号を裏す音源パルス列を入力して前記スペクトル パラメータとピッチパラメータと音源パルス列と を分離するデマルチプレクサ回路と、

ピッチパラメータと音源パルス列を用いて非一 機に分割された区間会体の音源信号を復元する音 海復元回路と、

復光された音源信号を用いて前記区間の音声信 号を合成する合成フィルタとを有する音声信号復 号化装置。

3 発明の詳細な説明

#### (産業上の利用分野)

本発明は音声信号を低いビットレートで効率的 に符号化、復号化するための音声信号符号化復号 化方法並びに音声信号符号化装置及び音声信号復 号化装置に関し、特に聴覚の特性にもとづいて音声 声を非一様に分割し、分割した区間において音声 信号の特徴を要すパラメータを求めて符号化、復 号化することのできる音声信号符号化復号化方法 並びにそれに用いる装置に関する。

#### 〔従来の技術〕

音声信号を低い伝送ビットレート、例えば8k b / s 以下で伝送する方法としては、8k b / s 程度ではビッチ予測マルチバルス符号化法、48 k b / s 程度ではビッチ補間マルチバルス符号化 法などが知られている。これらはいずれも音源信 号を複数個のパルスの組合せ(マルチパルス)で 褒し、声帯の特性をデジタルフィルタで表し、音 源パルスの情報とフィルタの係数を、一定時間区 間(フレーム)毎に求めて伝送している。この方 法の詳細については、前者は例えばOzawa、Araseki 氏による"Bigh Quality Multi-pulse Speech Coder with Pitch Prediction" (Proc. I.C.A.S.S.P. 講演番号33.3 1986) (文献1)に、後者については例えばOzawa、Araseki氏による 'low Bit Rate Multi-pulse Speech Coder with Natural Speech Quality" (Proc. I.C.A.S.S.P. 講演番号97, 1986) (文献2)に記載されている。これらの方法では、伝送情報量を抵減するために、音源パルス信号のピッチ予測やフレーム内の1つのピッチ区間に対してのみパルス列を求めることによって、伝送すべき音源パルス情報を低減している。

#### (発明が解決しようとする課題)

しかしながら、この従来の方法では、普敦パルス、フィルタ係数を求める区間長を一定(文献1.2では20ms)としていた。従って、母音区間ではほぼ周期的な波形が連続し音声の特徴が余り変化していないにも拘わらず、20msという短時間毎に情報を伝送するということで、非常に効率が悪く、低方、子音区間では違い音声の特徴の変化に追随出来ずに音質劣化が起こるという問題点が

あった。また、特にこの問題点はビットレートが 8 k b / s よりもかなり低い場合に顕著であった。 上述の問題を更に具体的に脱明すると、ます。 よく知られているように、母音区間は、発生速度 にも依存するが、一般に100~300mgと時間長が 長く、この半分以上は資戸信号の特徴が殆ど変化 しない定常区間とみなせる。更に、母音定常部で は、信号を零に即圧し情報を全く伝送しなくても、 音節明瞭度は殆ど劣化しないことが知られている。 低心 自然性は劣化する。従って、従来方法の如 く、これを短い20mm程度のフレーム区間毎に分 折して情報を伝送しているのでは効率が非常に悪 かった。一方、子音区間では音声の特徴の変化が 速いために、20m s のフレームでは長すぎて音声 の変化に対応した精度の良い分析が難しく、再生 音声の音質が劣化していた。

そこで、これらの問題点を改善するために、例 えばMarke!, Gray氏による"Linear Prediction of Speech"第10章(Springer-Verlag社, 1976) (文献3) にあるように、10ms程度の固定長フ レームで求めたスペクトルのフレーム間での差分の変化をもとに、フレーム長を固定区間の整数倍に可変にする方法が提案されているが、この改善策でも、問題がある。すなわち、かかる方法が受けては、上述のようにしたときに聴覚との対対の良くない特徴パラメータを用いてフレーム長の可変を行っていることと、フレーム長の可変をあると、音質が大きく劣化するという問題はがあった。

本発明の目的は、音楽信号伝送に必要な情報量を大幅に伝滅することができ、ピットレートを大 幅に下げても合成音声の聴覚的な劣化を非常に少 なくすることのできる音声信号符号化復号化方法 並びに音声信号符号化装置及び音声信号復号化装 置を提供することにある。

(課題を解決するための手段) 本発明の音声信号符号化復号化方法は、 離散的な音声信号を入力し聴覚の特性と対応の 良い方法により前記音声信号を非一様な区間に分割し、その分割された区間の全部または一部の区間における音源信号を複数個のパルス列の組合せで変して伝送し、

伝送されたパルス列を用いて前記区間の音源は 号を復元して前記音声は号を表す合成音声は号を 出力することを特徴としている。

また、本発明の音声信号符号化装置は、

入力した駆敗的な音声信号系列から聴覚の特性 と対応の良い特徴パラメータを摘出しそのパラメ ータを用いて音声活号を非一様な時間区間にセグ メンテーションするセグメンテーション回路と。

分割された音声信号の全部または一部の区間から短時間スペクトル特性を表すスペクトルパラメータとピッチパラメータとを計算するスペクトルパラメータ計算回路と、

分割された区間の会部または一部の区間における音源信号を表す複数個パルス列の組合せを計算 する音源パルス計算回路と、

スペクトルパラメータとピッチバラメークと音

更に、本発明の音声信号復号化装置は、

音声信号の短時間スペクトル特性を要すスペクトルパラメータとピッチパラメータと音楽信号を 支す音源パルス列を入力して前記スペクトルパラ メータとピッチパラメータと音源パルス列とを分 離するデマルチブレクサ回路と、

ピッチバラメータと音波パルス列を用いて非一様に分割された区間全体の音楽信号を復元する音 複復元両路と、

復元された音楽信号を用いて前記区間の音声信 号を合成する合成フィルタとを有することを特徴 としている。

#### (作用)

上記のようにして音声信号の符号化復号化を行 うため、音源信号伝送に必要な情報量を大幅に低 波でき、しかもピットレートを大幅に下げても合 成音声の聴覚的な劣化の非常に少ない符号化復号 化処理を行える。

音声信号符号化装置は、上記構成のセグメンデーション回路、スペクトルパラメータ計算回路。 音源パルス計算回路、マルチプレクサ回路を有す ることにより、上述のような符号化処理が行える。 音声信号復号化装置は、上記構成のデマルチプレ クサ回路、音源復元回路、合成フィルタを有する ことにより、その復号化処理を行える。

#### (実施例)

次に、本発明の実施例について図面を参照して説明する。

第1図は本発明による音声信号符号化復号化方法並びに音声信号符号化装置及び音声信号復号化装置の一実施例の構成を示すプロック図であり、また、第2図はその原理の説明に供するプロック図である。

本発明に係る音声信号の符号化復号化方法においては、送信側では離散的な音声信号を入力し聴覚の特性と対応の良い方法により音声信号を非一様な区間に分割し、分割された区間の全部または一部の区間における音源信号を複数個のパルス列

の組合せで表して伝送し、受信例では前記パルス 列を用いて前記区間の音源信号を復元して音声信 号を良好に表す合成音声信号を出力する。

以下、まず、本発明に従う符号化処理の原理について、第2図(a)を用いて規明する。図において、セグメンテーション尺度計算部400 は、音声信号を入力し、音声特徴変化の遠い子音部は高いであるような短時間区間(例えば5ms)毎に、聴覚の特性との対応の良いセグメンテーション尺度を計算する。ここではこの尺度は5ms毎に求めたしPCケブストラムC: (1≤1≤p)を用いて下式の様に書ける。

$$D(t) = \frac{1}{n} \sum_{i=1}^{n} a_{i}^{2} \cdots \cdots (1)$$

ここで、a.は、

$$\mathbf{a}_{ij} = \sum_{n=-n_0}^{n_0} \mathbf{C}_{ij}(a) + n / (\sum_{n=-n_0}^{n_0} n^{ij}) \qquad (2)$$

である。

尚、この計算法についての詳細な説明はFurui

氏による\* On the Role of Spectral Transition for Speech Peceptios\*と難した独文(J.Acoustical Society of America vol.80,pp 1016-1025,1986) (文献4) に記載されているので、ここでは詳細は省略する。また、(1)式の代わりにパワ項a。を含めた(3)式や他の良好な方法を用いることも出来る。

$$D (t) = \frac{a_0^2}{2} + \frac{p}{\frac{2n}{n-1}} a_1^2 + \cdots$$
 (3)

セグメンチーション部410 は、セグメンテーション尺度を入力して、音声信号を非一様に分割(セグメンチーション)する。これは前記(1)あるいは(3)式の尺度を用いて行う。まず前記尺度の極大値の付近年に音声信号をあらかじめ分割する。ここで、前記文献4に記されているように、前記尺度の極大値の前後数10m = の略分分にですから母音、母音から子音への調音結合部分にほぼ対応しており、音観知覚の際の聴覚的に非常に重要な部分であることが報告されている。従ってこのような聴覚的に重要な部分を除き前記尺度

がある程度連続的に小さくなる箇所で音声信号を セグメンチーションする。セグメンテーションし た様子を第2図(b)に示す。ここで第2図(b) の上段は音声波形、下段は動的尺度とセグメンテーションの一例を示す。

次に、1 P C, ビッチ分析部430 はセグメンデーションされた区間全体あるいはこの中の一な分の音声信号を分析してL P C 係数を求める。なシテーの音声信号を分析してL P C 係数を求める。なシテーションはは、セグメンテーンを表して、カーションは、カーションは、カーションは、カーションは、カーションは、カーションは、カーションは、セグメンテーションは、セグメンテーションは、ロロッチがあるができる。というないでは、セグルのでは、ロロッチがあるができないから、アーションは、ロロッチがある。とかできる。

音源計算部420 は、セグメンテーションされた 区間が母音定常部のときは、前記セグメンテーシ

ョン区間をピッチ区間の周期毎のサブフレームに分割し、そのうちの1つのピッチ区間について、 音源パルス列を計算する。ここで音源パルス列の 計算には、特願昭59-272435 号明細書(文献5) を参照することができる。

また、他のピッチ区間については、ピッチ区間 毎にピッチ区間の波形を良好に表すように接幅補 正接数を求める。

従って、本発明によれば、従来方式に比べビットレートを大幅に下げても「ピッチ区間の音源パルスの数を大幅に増やすことが可能であるため、 後述のように他のピッチ区間は振幅補正あるいは 補間処理を用いて復元するとしても、前記区間全 体の音楽信号を良好に要すことができる。

一方、前期セグメンテーション区間が母音定常 部でないときは、区間全体で音源パルス列を求め る。

送憶側の伝送情報は音源パルス列の提幅。位置、 セグメンテーションされた区間の長さを示すセグ メンテーション情報、ピッチ周期、判別情報、提 報補正係数である。受信側では、母音定常部の時は、伝送された音源パルス列の振幅と位置をピッチ周期毎に滑らかに変化させたり、セグメンテーションされた区間の間での音源信号に補間処理を推し、伝送されたピッチ区間以外のピッチ区間のパルス列を復元しセグメンテーションされた区間の音源信号を復元する。

次に、第1回を参照して説明する。

第1回において、送信側は音声信号符号化装置 を、また受信側は音声信号復号化装置をそれぞれ 含み、両者順には適宜の伝送路が設けられている。

音声信号符号化整置は、入力した慇懃的な音声 信号系列から聴覚の特性と対応の良い特徴パラメータを抽出し前記パラメータを用いて前記音声信 号を非一様な時間区間にセグメンテーションする セグメンテーション回路と、前記分割された音声 信号から短時間スペクトル特性を表すスペクトルパラメータ計算回路と、前記分割された区 フトルパラメータ計算回路と、前記分割された区 間の全部または一部の区間における音源信号を表 す複数個のパルス列の組合せを計算する音源パルス計算回路と、前記スペクトルパラメータと前記 ビッチバラメータと前記音源パルス列を組み合わせて出力するマルチプレクサ回路とを有する。

音声信号復号化装置は、音声信号の短時間スペクトル特性を要すスペクトルバラメータとピッチパラメータと音源信号を要す音源パルス列を入力して前記スペクトルバラメータと前記にカータと前記音調パルス列とを分離するデマルチプレクサ回路と、前記ピッチバラメークと前記音源パルス列を用いて非一様に分割された区間会体の音源信号を復元する音源復元回路と、前記復元された音源信号を用いて前記区間の音声信号を合成する合成フィルタとを有する。

音声信号符号化、復号化処理は、以下のように してなされる。

本発明の一実施例を示す第1図において、入力 端子500 から離散的な管庫信号を入力する。セグ メンテーション尺度計算回路505 は第2図(a) のセグメンテーション尺度計算部400 と同一の計 算を行い、セグメンテーション尺度を出力する。 セグメンテーション回路510 は第2図(a)のセ グメンチーション部410 と同一の処理を行い、音 声信号を非一様な区間にセグメンテーションし、 セグメンテーション区間の長さを要すセグメンテ --ション情報とセグメンテーションされた音声信 号を出力する。1PC、ピッチ計算回路520 は第 2 図 (a) のLPC、ピッチ分析部430 と同一の 処理を行い。セグメンテーションされた音声循号 について、LPC分析、ビッチ周期の計算及び、 セグメンテーションされた区間が母音定常部か否 かの判別を行い、1PC保敷、ビッチ周期、判別 情報を量子化器530 へ出力する。量子化器530 は これらの情報を所定のビット数で量子化しマルチ ブレクサ600 へ出力すると共に、これらを逆費子 化する。

重みづけ図路540 は、セグメンテーションされた音声信号と連載子化されたもPC係数を用いて前記信号に重みづけを施す。重みづけの方法は前記文献5の重みづけ回路(200) を参照することが

できる。インバルス応答計算回路560 は逆置子化されたLPC係数を用いてインバルス応答を計算の方法は前起文献5のインバルス応答計算の方法は前起文献5のインバルス応答計算回路570 は前記インバできる。自己相関関数を計算回路570 は前記インバルス応答の自己相関関数を計算回路(180) を参照することができる。相互相関関数計算回路550 は前記重みづけられた信号と前記インバルス応答との相互相関関数を計算にできる。相互相関関数を計算回路550 は前記重みづけられた信号と前記インバルス応答との相互相関関数を計算にできる。

音源パルス計算回路580 は、セグメンテーションされた区間が母音定常部の時は、前記第2図 (a) の裁明中で述べた様に、前記区間をピッチ 周期毎のサブフレームに分割して中央付近のサブフレーム区間について音源パルス列を計算する。また他のサブフレーム区間では削記第2図(a)

の説明中で述べたようにベルス列の復幅調正係数を各区間で1つずつ求める。一方、母音定常のでないときは、前記区間全体に対して音源ベルス列の計算法については前記文献5の器動信号計算回路(220)を参照することができる。量子化器590 は前記音源ベルスフの接幅,位置を所定のビット数で量子化してマルチプレクサ600 へ出力する。量子化器590 の動作は前記文献5の符号化回路(230)を参照することが前記文献5の符号化回路(230)を参照することが適いルス列、LPC係数、ピッチ周期、セグメンテーション情報、判別情報、振幅補正係数を組み合わせて出力する。

一方、受信帳では、デマルチプレクサ610 は、音なパルス情報。IPC係数、ピッチ周期、セグメンテーション情報、判別情報、提領博正係数を分離して出力する。音源パルス復号器620 は音源パルス列の提幅、位置を復号する。IPC ピッチ復号器640 はIPC係数、ピッチ周期を復号する。音源復元器830 は判別情報、セクメンテーシ

ョン情報を入力して、区間が母音定常部の時は、 復号した1ピッチ区間の音源パルス列を用いても グメンテーション区間全体の音源信号を復元し出 力する。ここで伝送されていないビッチ区間の音 タバルス列の復元法としては、ピッチ区間のバル ス全体をピッチ周期だけずらして位置を復元し、 漿幅に関しては擬幅補正係数を乗じて機幅を復元 する。この方法以外にも、隣接セグメンテーショ ン区間の音源パルス列を用いて排間処理によって 復元する方法などが知られており、この詳細につ いては前記文献5を参照することかできる。また これ以外にも他の周知な方法を用いることもでき る。一方、区間が母音定常部でないときには、受 信した音源パルス列を用いて前記区間全体の音源 信号を発生して出力する。補間器650 は復号した LPC係数、判別情報、ピッチ周期を用いて、セ グメンテーション区間が母音定常部のときはスペ クトル変化を清らかにするために、ピッチ周期毎 にLPC係数をPARCOR係数上で補間する。 - 方、前記区間が母音定常部でないときには係数

を辨別せずに合成フィルタ650 へ出力する。これは母音定常部以外では音声は号のスペクトル特徴の変化が速いので補簡によってかえって大きな歪が入ることを防ぐ為である。合成フィルタ660 はLPC保数、復元された音源信号、セグメンテーション信報を用いてセグメンテーション区間全体における音声信号を合成し端子670 を通して出力する。

トを大幅に下げても合成音声の聴覚的な劣化は非 常に少なく高い自然性が得られる。

上述した実施別はあくまで本発明の一実施例に 選ぎずその変形例を積々考えられる。

例えば、セグメンテーションされた区間が母音 定常部であるときには、相互相関関数計算回路550 は前配区間全体に対してではなく、前配区間の中 央付近の1ビッチ区間に対してのみ相互相関関数 を計算しても良い。これは実際に音源パルス列を 求めるのは1ビッチ区間であるためである。この 方法では特性は少し劣化するが資算量はほぼア/ N(ここでPはピッチ関類、Nは母音定常のセ グメンテーション区間の長さ)に低端できる。

また、音源パルスの計算法としては上述の実施 例の他に関知の良好な方法を用いることもできる。 これについては、2.0zawa \* A Study of Polse Search Algorithms for Multi-pulse Speech Codec Realization \*(J.Selected Area of Communications, pp. 1987) (文献 6) を参照することができる。 また、セグメンテーションされた区間が母音定常部のときは、音源パルス列を求める1ビッチ区間の位置としては、固定ではなく、最も良好な合成音声が得られるようなピッチ区間を探索して求めるようにすることもできる。この処理によって音質はさらに良好になるが演算量は若干増加する。具体的な方法については前記文献5を参照することができる。

また、合成フィルタ660 の係数の権間法としては、対数断面積出上や他のパラメータ上で補間することもできる。さらに補間法としては線形機間以外に対数補間等を用いることもできる。これらの方法の詳細については A.S. A tal 氏らによる "Speech Analysis and Synthesis by [tnear Prediction of the Speech Hava" (J Acoust. Soc. America, pp 637-855, 1971) (文献?) を参照することができる。

また、受信側でピッチ周期を拷問によって潜ら かに変化させることによって合成音質はさらに改 善される。

#### (発明の効果)

以上説明したように、本発明の音声信号符号化 復身化方法によれば、音声信号を符号化し伝送し て復号化したとき、音声信号を良好に表す合成音 声信号を得ることができる。従来の固定長フレー ムによるものや、あるいは固定長フレームで求め たスペクトルのフレーム間での整分の変化を基に フレーム長を可変にするものに比し實質の劣化を 少なくすることができる。聴覚の特性と対応づけ のよい特徴パラメータを用いて音声信号を非一様 に分割することができると共に、分割された区間 が音声の特徴の変化が殆どなく時間的にも長い母 音定常部のときはその区間のうち1つのピッチ区 間について音源パルス列を求め、母音定常部以外 のときは区間全体で音源パルスを求めることが可 能であり、音源信号伝送に必要な情報量を大幅に 低波することができると同時にビットレートを大 幅に下げても合成音声の聴覚的な劣化の非常に少 なく高い自然性が得られる符号化復号化処理を行 うことができるので、低いビットレートで効率的

に符号化、復号化する場合に適している。

さらに本発明によれば、音声俗号符号化復号化 方法を実施するのに好適な音声信号符号化装置及 び復号化装置が得られる。

#### 4. 図面の簡単な説明

第1図は本発明による音声信号符号化復号化法 法並びに音声信号符号化装置及び音声信号復号化 装置の一実施例の構成を示すプロック図。

第2図は本発明の説明に供する原理プロック図 及び被形図である。

- 400 セグメンテーション尺度計算部
- 410 セグメンテーション部
- 420 音源計算部
- 430 LPC ピッチ分析部
- 505 セクメンテーション尺度計算函路
- 510 ・・・セグメンテーション回路
- 520 1.PC ピッチ計算回路
- 530,590 登子化器
- 540 重みづけ回路
- 550 相互相関関数計算回路

580 インパルス応答計算回路

570 自己相關関数計算回路

600 マルチプレクサ

610 デマルチプレクサ

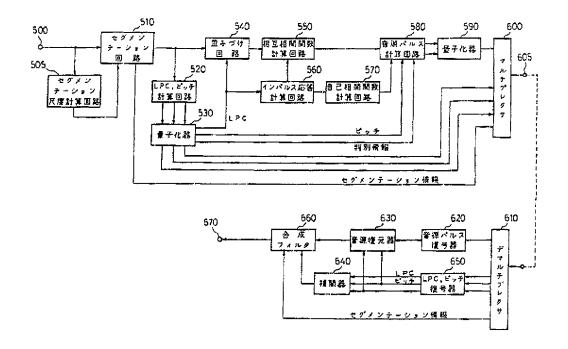
620.640 - - 復号器

630 金源復元器

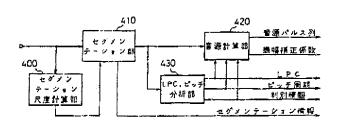
650 · 排間器

660 合成フィルタ

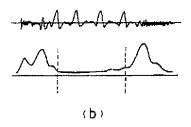
代理入 弁理士 岩 佐 森 幸



第 1 図



(a)



第 2 図